THE PENNSYLVANIA STATE UNIVERSITY SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

LIPID-MEDIATED LOCALIZATION OF POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE

ROHAN BASU SPRING 2017

A thesis submitted in partial fulfillment of the requirements for a baccalaureate degree in Biochemistry and Molecular Biology with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

David Boehr Professor of Chemistry Thesis Supervisor

Craig E. Cameron Eberly Chair in BMB Thesis Supervisor

Ming Tien Professor of Biochemistry Honors Adviser

Scott Selleck Department Head for Biochemistry and Molecular Biology

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

RNA viruses continue to exert a significant toll on human health globally. Poliovirus is a well-studied RNA virus and serves as an important model system for understanding the life cycles of such viruses. This research was focused on understanding the lipid binding interactions of the poliovirus RNA-dependent RNA polymerase (RdRp) to phosphatidylinositol-4-phosphate (PI4P). PI4P is thought to aid in the localization of poliovirus replication proteins, like RdRp, at host intracellular membranes to form viral "replication organelles." In order to interrogate the lipid-binding interactions of RdRp, nuclear magnetic resonance (NMR) spectra were collected on samples of poliovirus RdRp variants with and without PI4P to determine if and where any interactions between the lipid and protein occurred. The NMR results were also compared to previous molecular docking studies that suggested potential PI4P binding sites. Based on these NMR spectra, it appears that a docking site near the front side of the RdRp palm domain is favored. More work remains to be done in further interrogating these interactions in environments that more closely resemble biological systems.

TABLE OF CONTENTS

	iv
LIST OF TABLES	vi
ACKNOWLEDGEMENTS	vii
Chapter 1 : Introduction	1
Viral Replication Strategies	1
Picornaviridae/Poliovirus Renlication	1
Replication Organelle	2
Phosphotidylinositol_4_phosphate	2
Poliovirus RNA-dependent RNA Polymerase (RdRn 3D)	5
Poliovirus RdRn PI4P Binding Site	0 8
Nuclear Magnetic Resonance (NMR) and Chemical Exchange	0
Protein Labeling for NMR Study	10
Summary of Goals	12 14
Linid Nanodias	
FAPP1-PH as a Positive Control Methods Expression of FAPP1-PH	15 16 18 18
FAPP1-PH as a Positive Control Methods Expression of FAPP1-PH Purification of FAPP1-PH EAPP1 NMP. Semple Properation	15 16 18 18 18
FAPP1-PH as a Positive Control Methods Expression of FAPP1-PH Purification of FAPP1-PH FAPP1 NMR Sample Preparation	15 16 18 18 18 19
FAPP1-PH as a Positive Control Methods Expression of FAPP1-PH Purification of FAPP1-PH FAPP1 NMR Sample Preparation NMR Spectroscopy Results	15 16 18 18 18 19 19 19
FAPP1-PH as a Positive Control Methods Expression of FAPP1-PH Purification of FAPP1-PH FAPP1 NMR Sample Preparation NMR Spectroscopy Results Chapter 3: Evaluation of Interactions Between RdRp and PI Lipids	15 16 18 18 19 19 19 19 19
FAPP1-PH as a Positive Control Methods Expression of FAPP1-PH Purification of FAPP1-PH FAPP1 NMR Sample Preparation NMR Spectroscopy Results Chapter 3: Evaluation of Interactions Between RdRp and PI Lipids Poliovirus RdRp PI4P Binding Site	15 16 18 18 18 19 19 19 19 19 19 19 19
FAPP1-PH as a Positive Control Methods Expression of FAPP1-PH Purification of FAPP1-PH FAPP1 NMR Sample Preparation NMR Spectroscopy Results Chapter 3: Evaluation of Interactions Between RdRp and PI Lipids Poliovirus RdRp PI4P Binding Site Methods	15 16 18 18 19 19 19 19 19 19 19 19 19 16 16 16 18
 FAPP1-PH as a Positive Control	15 16 18 18 19 19 19 19 19 19 19 19 19 16 16 16 18
 FAPP1-PH as a Positive Control	15 16 18 18 19 19 19 19 19 19 19 19 19 16 16 16 18 18 18 19 19 19 19 19 19 19 19 19 19 19 19 19
 FAPP1-PH as a Positive Control	15 16 18 18 19 19 19 19 19 19 19 19 19 26 26 26 26 26 27 28
 FAPP1-PH as a Positive Control	15 16 18 18 19 19 19 19 19 19 19 19 19 19 19 19 19 19 19 19 19
 FAPP1-PH as a Positive Control	15 16 18 18 19 19 19 19 19 19 19 19 26 26 26 26 26 26 28 28 28 28 29 29 29
 FAPP1-PH as a Positive Control	

Docking Solutions 1 and 5	41
Creation of a New Docking Site	42
Future Directions	
Effective Modeling of Membrane Interactions	42
Alternative Approaches	43
In the Context of 3CD	43
BIBLIOGRAPHY	

LIST OF FIGURES

Figure 1. Outline of replication organelle formation.	4
Figure 2. General structure of phosphatidylinositol	6
Figure 3. Poliovirus 3D polymerase domains.	7
Figure 4. Proposed docking solutions of PI4P to the PV RdRp surface	9
Figure 5. Target docking solutions.	10
Figure 6. Chemical exchange signals	11
Figure 7. Isoleucines and methionines on 3D.	13
Figure 8. Biosynthetic precursor and terminally-labeled isoleucine	14
Figure 9. General structure of a lipid nanodisc.	15
Figure 10. FAPP1 Pleckstrin homology (PH) domain	17
Figure 11. Overexpression of FAPP1 PH domain.	20
Figure 12. 600 MHz HSQC of backbone labeled FAPP1-PH with and without PI4P	21
Figure 13. FAPP1-PH PI4P chemical shift changes.	22
Figure 14. 600 MHz HSQC of FAPP1-PH with and without PI4P nanodiscs	23
Figure 15. Absolute intensity change of selected residues.	24
Figure 16. Isoleucine-labeled spectrum of poliovirus 3D apoprotein	30
Figure 17. Isoleucine-labeled spectrum of the interaction of poliovirus 3D with PI4P	31
Figure 18. Isoleucine-labeled spectrum of the interaction of poliovirus 3D with PI4P lipid nanodiscs.	32
Figure 19. Sequence alignment of 3D with Poliovirus 3C binding domain	33
Figure 20. Docking site 6. View of poliovirus 3Dpol from the top.	33
Figure 21. Comparison of WT and M141I variant 3D	34
Figure 22. Top views of Poliovirus 3D	35
Figure 23. Supported lipid bilayer binding assays for 3D variants	36
Figure 24. Isoleucine-labeled spectrum of Poliovirus 3Dpol (D339A, S341A, D349A) with	n PI4P.37

Figure 25. Overlay of 3D variants.	38
Figure 26. Key protein residues in poliovirus 3D.	39
Figure 27. Crystal structure of 3CD and component 3C and 3D.	44
Figure 28. 3CD docking sites	45

LIST OF TABLES

Table 1. Ph	osphoinositide localization.	5
-------------	------------------------------	---

ACKNOWLEDGEMENTS

My undergraduate research career has been a maze of numerous and varying challenges, which I would not have been able to navigate were it not for a number of important people in my life. Thank you first to Dr. David Boehr for his continued support. David was always available to assist when I needed help, but gave me enough autonomy to grow as a scientist myself. He always challenged me and pushed me to do my best.

I would also like to thank my labmates Chris Bailey, Alyson Boehr, Becky D'Amico, Scott Gorman, Prasad Krishnan, Kate O'Rourke, Deb Sahu, and Jingjing Shi. I also would like to thank our collaborators in Dr. Craig Cameron's group and Dr. Nihal Altan-Bonnet's group.

Additionally, without the support of the Penn State BMB and Chemistry departments, I would never have any of these opportunities. I would also like to thank Angela Gronenborn, Elena Matei, and Mike Bergman for advising my work over the summers, which helped me to further develop my skills as a scientist.

I would also like to thank my friends and peers. Truly, without you I would not be the person I am today. You all inspire me and push me to be the best version of myself.

Finally, I would like to thank my family for their undying encouragement. There were times when I wanted to quit, but you never let me. You always reminded me about the good things the future will hold. Without my family, none of this would ever have been possible.

Chapter 1 :

Introduction

Viral Replication Strategies

By many estimates, viruses are the most abundant biological organisms on earth.¹ However, the term 'virus' is a catch-all term, as viruses show an extraordinary diversity of cellular life cycles.² Viruses with different life cycles also employ different survival strategies. Viruses store their genetic information in relatively small genomes. Many viruses hold a number of replication mechanisms in common, most of which underlie the ability to hijack host cell machinery, including membranes and proteins, to serve as factories for the production of viral components. This feature is perhaps the most amazing feature of viruses: they have the ability to take the hostile environment of the host cell and turn it into a functional home.

Two common classification systems exist for viruses: the Baltimore system and the International Committee of Taxonomy of Viruses system (ICTV).³ Proposed by David Baltimore in 1975, the Baltimore classification groups viruses into seven categories based on type of genome present and replication strategy. The groups include:

- I. double-stranded DNA (dsDNA)
- II. single-stranded DNA (ssDNA)
- III. dsRNA
- IV. positive sense (+) ssRNA
- V. negative sense (-) ssRNA

VI. ssRNA that replicate using a DNA intermediate: retrovirus (ssRNA-RT)VII. dsDNA that replicate using a ssRNA intermediate: retrovirus (dsDNA-RT)

The Baltimore system is perhaps a simpler means of classification than the ICTV system, which classifies viruses based on their properties, such as size, symmetry, type of nucleic acid used for the genome, and components of the virion. The ICTV system is taxonomical, and uses a hierarchical method of organization. The classification terms, ordered from least specific to most specific, are order (-virales), family (viridae), subfamily (virinae), genus (-virus), and species.⁴

Picornaviridae/Poliovirus Replication

Picornaviridae is an ICTV family that comprises a group of positive-sense ssRNA viruses broken into 35 genera including enteroviruses, hepatoviruses, and parechoviruses. While each different genus has distinctive characteristics, picornaviruses hold a remarkable number of qualities in common.⁵

Like all positive-sense viruses, picornaviruses both translate their positive-sense ssRNA genomes to produce proteins necessary for replication and transcribe their genomes to create new positive-sense ssRNA to serve as translation or replication templates.^{6,7}

Positive-sense RNA viruses comprise about a third of all known viruses and are important in a variety of roles and ecosystems, making them an interesting point of study.⁸ Picornaviridae contains a number of viruses that do not, as of yet, have cures or vaccines, designating it as relevant to biochemical research with biomedical applications.⁹ Because poliovirus is among the most well-understood picornaviruses, it was used as a model candidate for study.^{10,11}

Replication Organelle

What is perhaps most intriguing about the life cycle of positive-stranded RNA viruses is how they deal with the aforementioned issue of survival in a complex and unforgiving host cellular environment. To cope with the stress of host environments, these viruses form what have come to be known as

replication organelles.^{12,13} The viral replication organelle describes a compartment within the host cell in which the virus replicates. These replication organelles serve to localize factors and proteins necessary for viral replication, while also protecting the virus against host defenses.

Interestingly, viruses form these organelles by upregulating host cell phospholipid synthesis, hijacking host secretory pathways, and remodeling intracellular host membranes for their own use. In particular, the membranes of the Golgi apparatus and endoplasmic reticulum are utilized.^{14,15,16} Poliovirus specifically relies on modification of the Golgi membrane.¹⁷ In doing so, the virus is able to create a specialized microenvironment for replication of its own viral RNA.

Positive-sense viruses produce replication organelles of two morphotypes: membrane invaginations or host intracellular membranes and double-membrane vesicles (DMVs). Picornaviruses, along with coronaviruses and arteriviruses, produce DMVs, which morphologically resemble extrusions from the hijacked host intracellular membrane.¹⁸ In poliovirus, early stage single-membrane vesicles are gradually converted to DMVs that are 100-300 nm in diameter.

Single-membrane invagination replication organelles are generally thought to enclose the viral replicase complex and contain a small channel to the cytoplasm for the import of all necessary metabolites and export of nascent viral RNA. DMVs, however, do not contain these channels for transport, which has complicated explanations of different DMV viral replication mechanisms. Poliovirus is thought to replicate mostly on the cytoplasmic side of the single membrane replication organelle until the second membrane that forms the final DMV seals it.^{17,19} Poliovirus replication is thought to be most active in the early, single-membrane replication organelle stage.¹⁷ For other DMV viruses, it is unclear whether a similar mechanism occurs. Implicit in these explanations is the fact that import of metabolites and export of viral RNA must occur through this "sealed" membrane. The method for these transport mechanisms is still not understood, although the DMV is thought to protect the viral RNA against the host immune response.¹⁸

Initially, the virus likely depends on modified cellular metabolism proteins to activate the synthesis of phospholipids. Viral protein and replication complexes are then recruited due to the unique lipid composition of these membranes.¹⁶ In poliovirus, formation of this replication organelle is associated with an enrichment of phosphatidylinositol 4-phosphate (PI4P) moieties. ADP-ribosylation factor (Arf1) is activated by a guanine exchange factor (GEF). The active Arf1-GTP binds the membrane and leads to the recruitment of PI4KIIIβ kinases, which enrich in the membrane in PI4P (Figure 1).



Figure 1. Outline of replication organelle formation. During poliovirus infection, replication organelles (shown as red ovals) develop on the trans-golgi network. PI and PIP kinases enrich the membrane in specific phospholipids, which are thought to aid in the localization of viral replication proteins.

Phosphotidylinositol-4-phosphate

Phosphoinositides are a noteworthy class of lipids that are enriched at these replication organelles during the membrane remodeling process and have been implicated in the recruitment of viral replication machinery. Different PI moieties show enrichment, at different membranes, and as such recruit different proteins to different locations within the host cell, as shown in Table 1.²⁰

Phosphoinositide	Cellular Location
PI3P	Early endosome,
	Multivesicular bodies
PI4P	Golgi, Vesicles
PI(3,4)P ₂	Early endosome, Vesicles
PI(3,5)P ₂	Late endosome, Lysosome,
	Multivesicular bodies
PI(4,5)P ₂	Plasma membrane
PI(3,4,5)P ₃	Plasma membrane

Table 1. Phosphoinositide localization. Adapted from Yan Mei Chan.^{15,21,22}

Of particular interest is phophotidylinositol-4-phosphate (PI4P), which has been linked with the recruitment of viral proteins, including viral RNA polymerases.²³,²⁴



Figure 2. General structure of phosphatidylinositol. Phosphatidylinositols can be phosphorylated at any of the positions inositol ring nonadjacent to the lipid tail position. In the case of PI4P, the position marked with an asterisk is phosphorylated.

For the purposes of this study, free, water-soluble short-chain PI4P was added to NMR solutions. It is important to note that *in vivo* it is unlikely that short-chain PI4P will be the form of phospholipid interacting with the protein. Rather, it is more likely the PI4P tail is embedded in the intracellular membrane of interest.

Poliovirus RNA-dependent RNA Polymerase (RdRp, 3D)

Poliovirus RdRp (3D) has been thought to bind preferentially to PI over other phospholipids; in particular, 3D binds PI4P and PI(4,5)P₂.¹⁵ However, 3D does not contain the traditional PI4P-binding pleckstrin homology (PH) domain. As such, it may be using a different recognition motif, which if identified and characterized, could have implications for both viral and host proteins. PI4P was shown to bind the poliovirus RNA-dependent RNA polymerase (3D) and the viral protease (3C) through membrane strip assays. Dr. Yan Mei Chan's work showed the location of binding with 3C to be on an N-terminal α helix, but the exact binding site on 3D is not yet known.²¹ The poliovirus RNA-dependent RNA polymerase is a 53 kDa, 461 amino acid protein that structurally resembles a cupped right hand shape, with helical finger and thumb domains linked by an N-terminal region.^{25,26}



Figure 3. Poliovirus 3D polymerase domains. Front view of the poliovirus RdRp (PDB 1RA6).

The template RNA enters the catalytic site on the palm through a tunnel created by the fingertip region. Binding and localization of the substrate nucleotide triphosphates (NTP) require two divalent metal cations and are mediated by highly conserved aspartate residues in the catalytic site.^{27,28,29} This binding cleft is about 15 Å wide and is able to fit A-form RNA duplexes. Uridylylation of the viral

protein VPg serves as a primer to initiate transcription, wherein a tyrosine residue is positioned such that transcription can be initiated from the hydroxyl group.³⁰ The polymerization reaction proceeds by way of the metal cations. The first cation activates the 3'-OH necessary for nucleophilic attack, while the second cation positions the NTP properly for elongation. The protein is also thought to cooperatively bind substrates, and is prone to forming homodimers.³¹

Poliovirus RdRp PI4P Binding Site

Poliovirus 3D is known to bind PI4P, as confirmed by lipid binding assays.³² However, the exact nature and location of this binding interaction is still unknown. The purpose of this study was to determine where and how 3D binds PI4P. Previous work and sequence homology with poliovirus 3C provide preliminary suggestions as to where this interaction may be occurring.²¹ PI4P contains negatively charged headgroup, which means it has the potential to interact with positively charged regions of 3D. Using molecular docking, a few PI4P docking sites on PV 3D were previously proposed.^{33,34,35}



Figure 4. Proposed docking solutions of PI4P to the PV RdRp surface. The docking solutions show the PI4P molecule in the proposed docking sites in the conformation proposed. These simulations were performed using Autodock.

However, a few of these docking solutions put the PI4P in a conformation that would likely not exist at the membrane, as the lipid tails of the PI4P are likely membrane bound, with the charged head group in the cytosol. As such, it is important to note that these simulations are artificial. Nonetheless, these docking simulations pose interesting questions as to where exactly PI4P is binding to 3D *in vivo*.



Figure 5. Target docking solutions. The three proposed PI4P docking solutions selected for study are shown. Residues important in the interactions are labeled, and the PI4P is shown in its proposed binding conformation.

Nuclear Magnetic Resonance (NMR) and Chemical Exchange

NMR was used to interrogate these potential interactions. In particular, ligand titration experiments were used, where a spectrum was taken first of the apoprotein and then of the protein-ligand mixture. The goal of such experiments is to identify changes in signal chemical shifts and intensities, which can then be correlated with specific amino acid residues and may indicate protein-ligand interactions. NMR is also able to give information on the chemical exchange of the protein-ligand interaction.

Chemical exchange is a term used to describe changes in chemical environment. In an NMR experiment, the existence of chemical exchange means the NMR probe is seeing two or more different chemical environments at different points in time. In the slow exchange regime, the nuclei move slowly between these different states so separate signals are seen for each state with unique chemical shifts and peak intensities (Figure 6). In the fast exchange regime, the two states interconvert rapidly enough that the NMR spectrum shows a population-weighted average signal of the different states (Figure 6).³⁶ Noting changes in chemical shifts and peak intensities can thus provide insight into the kinetics and thermodynamics of protein-ligand interactions.



Figure 6. Chemical exchange signals. Figure from Denk, M.³⁶ In the fast exchange regime, an average of the two states is seen, where as in slow exchange, both unique states show distinct signals.

Protein Labeling for NMR Study

Due to the size constraints of solution-state NMR, protein NMR requires some ingenuity, especially when studying larger proteins. The most common solution to issues of size in protein NMR is the use of selective isotopic labeling. In particular, specific residues can be labeled to serve as molecular "probes" of the overall protein structure. Imperative to choosing these probes is the selection of residues with high levels of coverage. That is to say, residues selected for labeling should be high in number in the primary amino acid sequence of the protein, such that enough probes exist to collect meaningful data. Additionally, an important consideration for good coverage of the protein is the distribution of the selected residues. These residues should be well dispersed throughout the protein structure, such that all portions of the structure can be probed.^{37,38} The most important consideration for this study is choosing residues that can act as probes in or near the PI4P binding regions predicted by the molecular docking studies. For the purpose of this study, the main residue selected was isoleucine, as there are 25 well-dispersed isoleucines throughout the 461 amino acid sequence of 3D (Figure 7) and there are a number of isoleucines near the proposed docking sites. For one part of the study, methionine was also used, as there are 17 well-dispersed methionines in the primary amino acid sequence of the protein, a number of which are also near the proposed docking sites (Figure 7).^{39,40,41}



Figure 7. Isoleucines and methionines on 3D. A cartoon rendering of poliovirus 3D is shown with the isoleucine residues highlighted as red spheres (A) and the methionines highlighted as magenta spheres (B).

Isoleucine labeling of the protein can be achieved by adding a labeled biosynthetic precursor of isoleucine, α -ketobutyric acid, to the bacterial culture, which the bacteria then incorporate into the induced protein (Figure 8). Methionine is easily labeled by simply adding ¹³C-labeled methionine to the minimal growth media, in lieu of ¹²C-methionine.

Many of these isoleucines are near the proposed docking sites, and as such can be used as probes for interactions at or near those sites. I16, I17, I58, and I176 are found near docking site 1, I397, I401, I436, and I441 are found near docking site 5, and I16, I17, I130, and I176 are found near docking site 6. Similarly, M6 is near docking site 1, M392 and M394 are near docking site 5, and M101, M123, M141, and M286 are near docking site 6.

A number of these residues are also near key structural RdRp motifs. M323, I324, and I331 are near the GDD motif, structural motif C, which coordinates metal ions necessary for activity.²⁸ I67 neighbors K66, a residue crucial for enzyme activity in its role of interacting with the incoming nucleotide triphosphate (NTP). I295, M299, and I300 surround N297, which interacts with the incoming nucleotide 2'-OH. Additionally, a number of the primary structure isoleucines and methionines are near proposed PI4P docking sites. As such, by monitoring the behavior of these residues with and without PI4P can help shed light on the protein-phospholipid interactions.



Figure 8. Biosynthetic precursor and terminally-labeled isoleucine. Terminal methyl ¹³C α -ketobutyric acid (2-ketobutyric acid) is added as a biosynthetic precursor to the bacterial culture shortly before expression. The bacteria use the labeled precursor to incorporate isoleucines with terminal methyl groups labeled with ¹³C into the expressed protein. The labeled carbon is circled in red.

In using these amino acid probes, a cleaner picture of the overall protein can be obtained. Were

all the residues to be labeled in a similar manner, there would simply be too many signals to interpret.

Summary of Goals

The overall goals of this study were to first validate the utility of solution-state NMR in

characterizing protein-lipid interactions for titration studies. After doing so, solution-state NMR was used

to probe interactions between poliovirus 3D protein and PI4P to interrogate the binding site for lipid-

mediated localization of 3D to intracellular host membranes.

Chapter 2 : Testing PI4P-containing Nanodiscs with the Model PI4P-binding Protein FAPP1

Lipid Nanodiscs

The major goal of the study was to validate or refute the 3D-PI4P docking studies using solutionstate NMR. While PI4P has been shown to bind 3C and 3D, these assays were performed using shortchain, free PI4P. As discussed earlier, it is unlikely that free PI4P is interacting with the 3D protein *in vivo*. As such, lipid nanodiscs were used to serve as a sort of membrane surrogate.

Lipid nanodiscs are an interesting tool for testing interactions in a more membrane-like environment. They are composed of a phospholipid bilayer wrapped in a protein scaffold for structure and are meant to model cellular membranes (Figure 9).



Figure 9. General structure of a lipid nanodisc. The phospholipids, shown with orange head groups and tan-colored tails are held together by structural scaffold proteins, shown in blue and green. Figure from Mitra, N.⁴³

The composition of phospholipids can be altered to elucidate specific lipid binding interactions or to better model specific types of membranes. In essence, the nanodiscs present the lipids and membrane components as they may appear in a real membrane, rather than as free-floating molecules. For the purpose of this study, nanodiscs enriched in PI4P were used to better mimic intracellular Golgi membranes.⁴³

While lipid nanodiscs can be a useful tool for the study of membrane interactions, they are limited by issues in modeling natural membrane fluidity, orientation, and curvature. Also inherently limiting in the nanodisc simulations is the fact that the concentration of PI4P is significantly less in a nanodisc than when free short-chain PI4P is used. Furthermore, the exact concentration of PI4P in the nanodiscs is difficult to pin down, but it is roughly on the order of 500 times lower than the 2.7 mM used for the free short-chain PI4P, as calculated from mole fractions of the nanodisc synthesis.⁴⁴ Nonetheless, for NMR study, lipid nanodiscs serve as valuable tools in modeling protein-lipid interactions in a manner that more closely resembles the interaction *in vivo*.

FAPP1-PH as a Positive Control

Four-phosphate adaptor protein (FAPP1) is a regulator of secretory transport from the trans-Golgi network to the plasma membrane and is a well-studied binder of PI4P. FAPP1 binds PI4P through a 110 residue N-terminal pleckstrin homology (PH) domain.⁴⁵ PH domain is the quintessential PI binding domain, allowing FAPP1 to serve as a positive control for PI4P binding assays.⁴⁶



Figure 10. FAPP1 Pleckstrin homology (PH) domain. Structure of FAPP1 protein is shown, with residues in the PH domain labeled. A ball and stick rendering of PI4P is overlaid in the binding site. Figure from Prestegard, et al.⁴⁴

Given that FAPP1 has a known PI4P-binding domain, it emerged as an attractive candidate for confirming the validity of the NMR PI4P titration experiments. Although it is not necessary that 3D contains a pleckstrin homology domain like FAPP1, the FAPP1 assay is useful in determining the functionality of the free PI4P, nanodiscs, and the NMR study overall.

For this study, only the PH domain of FAPP1 was purified. Key in the PI4P binding interaction are basic residues K7, R18, K41, and K45. Binding with short-chain PI4P was shown in the literature to be in the millimolar range and in the fast-exchange regime.⁴⁵

The solution-state NMR experiments were verified in their ability to observe changes upon adding PI4P to a protein solution. The nanodiscs also showed changes in the spectra upon addition to the protein solution, but there appeared to be size-constraint issues that may be problematic for studying a 3D-nanodisc complex.

Methods

Expression of FAPP1-PH

The expression pSUMO construct was transformed into *E. coli* BL21 (DE3) cells and plated on LB-agar plates with kanamycin. The next morning 4x10 mL M9 media with kanamycin in 50 mL Falcon tubes were inoculated with single colonies. The OD was measured at 600 nm at 8 and 12 hours. The culture with the highest OD₆₀₀ at 12 hours/OD₆₀₀ at 8 hours ratio was streaked onto M9 agar plates in 100% D2O with kanamycin. The plates were incubated for 20-22 hours at 37°C, after which 4x10 mL M9/50% D₂O cultures were inoculated. The OD₆₀₀ was taken at 16, 20, and 24 hours. A 1 mL aliquot of the culture with the highest OD₆₀₀ at 24 hours/OD₆₀₀ at 20 hours ratio was used to inoculate 50 mL M9 in 100% D₂O and shaken overnight at 37°C and 250 rpm. This culture was then used to inoculate 950 mL of M9 in 100% D₂O with ¹³C-glucose and ¹⁵NH₄Cl. This culture was then shaken at 37°C and 250 rpm until the OD₆₀₀ = 0.6. 0.25 g IPTG in 5 mL D₂O was sterile filtered into the culture, and the culture was induced at 25°C for 20-24 hours. Finally, the cells were harvested at 10,000 x g for 30 minutes.

Purification of FAPP1-PH

First, base buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl), resuspension buffer (base buffer + 10 mM imidazole, 1mM PMSF, 1.4 ug/mL Pepstatin A, and 1 ug/mL Leupeptin), equilibration buffer (base buffer + 10 mM imidazole), wash 1 buffer (base buffer + 10 mM imidazole), wash 2 buffer (base buffer + 50 mM imidazole), and elution buffer (base buffer + 500 mM imidazole) were prepared. The cell pellets were then resuspended in 5 mL resuspension buffer per 1 gram of cell pellet and lysed using a French press. The lysate was then centrifuged at 16,500 rpm for 30 minutes at 4°C. The supernatant was run through a Ni-NTA column equilibrated in equilibration buffer. The column was washed with 10 column volumes of wash buffer 1 and 5 column volumes of wash buffer 2. The protein was then eluted using 10 column volumes of elution buffer.

FAPP1 NMR Sample Preparation

The protein was spin concentrated to about 150 uM and exchanged into NMR buffer using Thermofisher zeba spin desalting columns. The NMR buffer contained 10 mM potassium phosphate, 200 mM NaCl, 10 mM K₂SO₄, 2 mM MgCl₂, 5 mM DTT in 5% D₂O.

NMR Spectroscopy

Heteronuclear single quantum coherence (HSQC) experiments were used to study the FAPP1-PH to PI4P interactions. HSQC is a fundamental two-dimensional NMR experiment. Magnetization is transferred from a hydrogen to the attached ¹⁵N nucleus through J-coupling, and then transferred back to the hydrogen, wherein it is detected by the NMR instrument. HSQC is used to show all labeled H-N correlations, which are usually on backbone amides of the protein chain. These NMR experiments were performed on a Bruker Avance III 600 MHz spectrometer.⁴⁷

Results

The expression of the FAPP1 protein, performed as detailed in the methods section (Figure 11) shows the elution of the FAPP1-PH domain at 12.2 kDa, as expected.



Figure 11. Overexpression of FAPP1 PH domain. The protein eluted at 12.2 kDa. The purification was performed as described in the methods section. An 8 mL Ni-NTA column was used to purify the Histagged protein. The column was washed with a sodium phosphate buffer (pH 8) and eluted with increasing imidazole concentrations. Wash 2 used 50 mM imidazole, where as the elution buffer used 500 mM imidazole.

Following the overexpression and purification of the FAPP1 protein, NMR spectra were

collected. The initial FAPP1 experiment was performed by collecting spectra of the apoprotein and then

of the protein with PI4P. These experiments were conducted in order to verify the functionality of the

experiment, the protein, and the PI4P ligand (Figure 12).



Figure 12. 600 MHz HSQC of backbone labeled FAPP1-PH with and without PI4P. The apo-state FAPP1-PH (150 uM) is shown in black. Shown in red is the protein with 2.7 mM short-chain PI4P added. The buffer conditions are 10 mM potassium phosphate, 200 mM NaCl, 10 mM K₂SO₄, 2 mM MgCl₂, 5 mM DTT in 5% D₂O.

Given that only the PH domain was purified, it was expected that the basic lysine and histidine residues in the binding pocket vital for binding would show chemical shift changes. In particular, K7, W8, T9, L12, W15, Y29, K41, G42, I44, K45, E50, I65, Q69, H70, and Y72 were previously shown to have the most pronounced chemical shift changes upon PI4P binding.⁴⁵ Because obvious chemical shift changes are seen in a number of these residues, this data suggests an interaction. These binding

21

interactions are likely in the fast-exchange regime, as noticeable chemical shift changes are seen with peaks unique to the bound and unbound states. In particular, large chemical shift changes are seen in the signals corresponding to V4, K52, H54, and T100 (Figure 13).



Figure 13. FAPP1-PH PI4P chemical shift changes. Cartoon of FAPP1-PH protein with PI4P-binding pocket residues shown as red spheres. Residues corresponding to noticeably perturbed peaks in Figure 12 are shown as blue spheres. PDB ID: 3RCP.

Interestingly, the residues shown in blue (Figure 13) are not all near the known PI4P binding site.

The observed chemical shift changes in these residues may be due to conformational changes of the

protein upon ligand binding. Once it was verified that NMR could be used to observe changes upon

adding PI4P to a protein solution, a test experiment using PI4P nanodiscs was performed to validate the

nanodisc assay (Figure 14).



Figure 14. 600 MHz HSQC of FAPP1-PH with and without PI4P nanodiscs. The apo-state FAPP1-PH (150 uM) is shown in black. Shown in red is the protein with PI4P nanodiscs added. The buffer conditions are 10 mM potassium phosphate, 200 mM NaCl, 10 mM K₂SO₄, 2 mM MgCl₂, 5 mM DTT in 5% D₂O.

The nanodisc experiments yielded an interesting result that would inform the work going forward. The first point to note is that there were no apparent chemical shift changes. However, the NMR spectrum in the presence of the nanodisc (shown in blue in Figure 14) is much less intense than the NMR spectrum in the absence of the nanodisc (shown in black in Figure 14), suggesting that the protein and nanodisc are in slow exchange. The extra noise seen in the presence of the nanodisc is also likely due to the low intensity of the resonances (Figure 15). This decrease in intensity may provide an initial indication of a larger complex between FAPP1-PH and the lipid nanodiscs that may be too large for NMR detection. That is, the nanodisc-protein complex may be forming, but the corresponding resonances for the complex are not observable. However, the effects of the interaction are indicated directly by the loss of peak intensity in the apoprotein resonances.



Figure 15. Absolute intensity change of selected residues. Peak intensities of residues visible and resolvable on both spectra in Figure 14 are highlighted. Many of the signals seen in the apoprotein spectrum were not seen in the nanodisc spectrum. Many of the other signals were too close together to be resolved with the given data. Residues shown here were visible on both spectra and were resolvable.

Because there were no easily observable chemical shift changes, the intensity changes of signals with and without ligand were analyzed. The peak intensity changes between the apoprotein spectrum and the nanodisc spectrum (Figure 14) were surprisingly uniform, even for some of the residues that showed chemical shift changes in Figure 13, such as V4 and T100. However, a smaller percent signal decrease is seen for H54 upon addition of the nanodiscs and the K52 peak disappears all together. This result suggests that the protein-nanodisc complex is forming, and that the nanodisc interactions are leading to additional relaxation processes for the K52 and H54 resonances. Altogether, the results suggest that the

protein-nanodisc complex is forming, but that the resonances corresponding to the complex are not observable, perhaps because the protein-nanodisc complex is outside the size constraints of solution state NMR. This finding presents problems if we are to use the lipid nanodiscs to study the 3D interaction, since the 3D is even larger than the FAPP1-PH protein.

Chapter 3: Evaluation of Interactions Between RdRp and PI Lipids

Poliovirus RdRp PI4P Binding Site

Though poliovirus 3D is known to bind PI4P, the specifics of this binding interaction are still unknown. This part of the study was conducted to determine where and how this binding interaction is occurring. The main goal was to evaluate previous docking studies using solution-state NMR using specific methionine and isoleucine probes that can interrogate these proposed binding sites. The solutionstate NMR experiments appeared to be consistent with PI4P interactions at/near docking sites 1 and 5. However, there were anti-dimerization changes in the "wildtype" 3D construct (L446D, R455D) that were necessary for the production of NMR samples at high protein concentrations. These residue changes are located in docking site 5. As such, a different anti-dimerization variant (D339A, S341A, D349A) was studied. However, the study of the second anti-dimerization variant did not resolve the issues raised with the first variant, as a chemical shift change was only seen for I316, which was not near any of the proposed docking sites. Upon compiling all of the data, it appeared that there may be a multivalent binding surface that 3D uses to interact with PI4P moieties at intracellular host membranes.

Methods

Expression of ¹³C-Methionine Labeled Poliovirus RdRp via Auto-induction

The poliovirus 3Dpol mutant was expressed using an auto-induction protocol. First, 25 ng of plasmid was transformed into 100 uL B834* pRARE competent cells and grown at 37°C on kanamycin/chloramphenicol LB agar plates overnight. A smear of colonies was used to inoculate the pH

7.6 starting media (88.58 mL ddH₂O, 5 mL 20X NPS, 2 mL 17 amino acid mix (10mg/mL), 1.5 mL ¹³C-labeled methionine (10 mg/mL), 2 mL B12 (5 mM), 0.125 dextrose (40%), 0.1 mL 1M MgSO₄, 0.1 mL 1000X trace metals mix, 0.3 mL kanamycin (25 mg/mL), 0.3 mL chloramphenicol (20 mg/mL)) the next day, which was shaken at 37°C, 250 rpm until it reached OD₆₀₀=1.0. The starting media was then used to inoculate the auto-induction media (433.5 mL ddH₂O, 25 mL 20X NPS, 10 mL 50X 5052, 10 mL 17 amino acid mix (10mg/mL), 7.5 mL ¹³C-labeled methionine (10 mg/mL), 10 mL B12 (5 mM), 0.125 dextrose (40%), 0.5 mL 1M MgSO₄, 0.5 mL 1000X trace metals mix, 1.5 mL kanamycin (25 mg/mL), 1.5 mL chloramphenicol (20 mg/mL). This culture was allowed to grow at 37°C and 250 rpm until it reached OD₆₀₀=0.3-0.5. The temperature was then changed to 25°C and the culture allowed to grow overnight. The cells were then harvested at 4000 rpm, 4°C for 30 minutes, washed (10 mM Tris, pH 8.0, 1 mM EDTA), and spun again at the same conditions.

Expression of ¹³C-Isoleucine Labeled Poliovirus RdRp via IPTG-induction

The expression pSUMO construct was transformed into *E. coli* BL21 (DE3) cells with pRARE plasmid and plated on LB-agar plates with kanamycin and chloramphenicol. The next morning 4x10 mL M9 media with kanamycin and chloramphenicol in 50 mL Falcon tubes were inoculated with single colonies. The OD was measured at 600 nm at 8 and 12 hours. The culture with the highest OD₆₀₀ at 12 hours/OD₆₀₀ at 8 hours ratio was streaked onto M9 agar plates in 100% D2O with kanamycin and chloramphenicol. The plates were incubated for 20-22 hours at 37°C, after which 4x10 mL M9/50% D₂O cultures were inoculated. The OD₆₀₀ was taken at 16, 20, and 24 hours. 1 mL of the culture with the highest OD₆₀₀ at 20 hours ratio was used to inoculate 50 mL M9 in 100% D₂O and shaken overnight at 37°C and 250 rpm. This culture was then used to inoculate 950 mL of M9 in 100% D₂O. This culture was then shaken at 37°C and 250 rpm until the OD₆₀₀ = 0.6, at which point 70 mg α -ketobutyric acid in 5 mL D₂O was sterile filtered into the culture. This culture was shaken for an

additional hour. Next, 0.25 g IPTG in 5 mL D_2O was sterile filtered into the culture, and the culture was induced at 25°C for 20-24 hours. Finally, the cells were harvested at 10,000 x g for 30 minutes.

Purification of Poliovirus RdRp

First, buffer B (100 mM potassium phosphate pH 8.0, 500 mM NaCl, 5 mM imidazole, 5 mM βmercaptoethanol, 60 uM ZnCl₂, and 20% w/v glycerol) and buffer C (100 mM potassium phosphate pH 8.0, 500 mM NaCl, 5 mM β-mercaptoethanol, 60 uM ZnCl₂, and 20% w/v glycerol) were prepared. The cell pellets were then resuspended in 50 mL lysis buffer (50 mL buffer B, 1.4 ug/mL pepstatin A, 1 ug/mL leupeptin, 1 mM PMSF, and 0.1% N-P40) and lysed through sonication. The lysate was then centrifuged at 30,000 x g for 30 minutes at 4°C. The supernatant was run through a Ni-NTA column equilibrated in buffer C1 (buffer C, 5 mM imidazole). The column was washed with buffer C1, buffer C2 (buffer C, 5 mM imidazole). The protein was then eluted using high imidazole concentrations in buffer C3 (buffer C, 50 mM imidazole) and buffer C4 (buffer C, 500 mM imidazole). The His tag and SUMO domain were cleaved from 3D using Ulp1 protease. The protein was then dialyzed in 80 mM Tris-HCl pH 8.0, 500 mM NaCl, 20% w/v glycerol, 10 mM β-mercaptoethanol, and 60 uM ZnCl₂ overnight. The resulting protein solution was then dialyzed in 100 mM potassium phosphate pH 8.0, 20% w/v glycerol, 10 mM β -mercaptoethanol, and 60 uM ZnCl₂ for 2-3 hours. The protein was then run through another Ni-NTA column using an identical procedure as before. Finally, the protein was run over a Q sepharose column at 0.5 mL/min. The column was washed with 10 column volumes of buffer F (50 mM Tris pH 8.0, 25 mM NaCl, 1 mM DTT, 20% glycerol). Finally, the protein was eluted with 4 column volumes of buffer F (50 mM Tris pH 8.0, 500 mM NaCl, 1 mM DTT, 20% glycerol).

Poliovirus RdRp NMR Sample Preparation

The protein was spin concentrated to about 150 uM and exchanged into NMR buffer using Thermofisher zeba spin desalting columns. Two possible NMR buffers were used. The first contained 25 mM potassium phosphate, 150 mM NaCl, 1 mM DTT, 0.02% NaN₃ in 100% D₂O. The second contained 20 mM HEPES (pH 8.0), 0.02% sodium azide, 150 mM NaCl, 1 mM DTT in D₂O.

NMR Spectroscopy

For the 3D NMR experiments, NMR HMQC (heteronuclear multiple quantum coherence) experiments were conducted. HMQC is very similar to HSQC, except that in HMQC both the proton and the nitrogen are allowed to evolve/signal. These NMR experiments were performed on Bruker Avance III 500 and 600 MHz spectrometers.

Results

An NMR HMQC experiment was run on the NMR 3D-wildtype (WT). It is important to note that polymerization of 3D proteins is prohibitive for NMR study. Therefore, the NMR 3D WT used was actually 3D variant L446D, R455D. These residues were changed to prevent polymerization 3D. The results of the Isoleucine-labeled NMR WT are shown in Figure 16:



Figure 16. Isoleucine-labeled spectrum of poliovirus 3D apoprotein. 600 MHz HMQC spectrum of the poliovirus 3Dpol apoprotein (300 uM), in black. The buffer conditions were 25 mM potassium phosphate, pH 8.0, 150 mM NaCl, 1 mM DTT, and 0.02% sodium azide in D₂O.

Notable residues near the potential binding sites include I58 (near docking site 1), I397 and I436

(near docking site 5), and I130 (near docking site 6). With the baseline wildtype spectrum established,

PI4P was added to create a basis for the interaction of PI4P with the wildtype protein (Figure 17).



Figure 17. Isoleucine-labeled spectrum of the interaction of poliovirus 3D with PI4P. Superimposed 600 MHz HMQC spectra of poliovirus 3D (300 uM) unbound, in black, and poliovirus 3D in the presence of 8.1 mM PI4P, in red. The buffer conditions were 25 mM potassium phosphate, pH 8.0, 150 mM NaCl, 1 mM DTT, and 0.02% sodium azide in D₂O.

Noticeable chemical shift changes are highlighted with the associated residue number. In

particular, there are small chemical shift changes associated with I397 and I436, which are near docking site 5. I58, associated with docking site 1, also showed a substantial chemical shift change, suggesting that docking site 1 may also be undergoing some sort of change upon the addition of PI4P. As a basis of comparison, the free PI4P spectrum was then compared with a PI4P lipid nanodisc spectrum (Figure 18).



Figure 18. Isoleucine-labeled spectrum of the interaction of poliovirus 3D with PI4P lipid nanodiscs. Superimposed 600 MHz HMQC spectra of poliovirus 3D (300 uM) unbound, in black, and poliovirus 3D in the presence of PI4P lipid nanodiscs, in red. The buffer conditions were 25 mM potassium phosphate, pH 8.0, 150 mM NaCl, 1 mM DTT, and 0.02% sodium azide in D₂O. The PI4P concentration is about 6 uM.

Unfortunately, there were no observable chemical shift or peak intensity changes induced upon adding PI4P-containing nanodiscs to 3D (Figure 18). One of the challenges of this experiment is that the concentration of PI4P is on the order of 500 times lower than that of the free, short-chain PI4P added in Figure 18 (~6 uM). Under these conditions, binding interactions between 3D and the nanodiscs may be too weak to cause any changes in the NMR spectrum. Due to the work of collaborators in the Altan-Bonnet group, docking site 6 was slated for further study. This site was favored based on sequence similarity with the PI4P binding domain of 3C (Figure 19).



Figure 19. Sequence alignment of 3D with Poliovirus 3C binding domain. Conserved residues are denoted with (*), conservative mutations with (:), and semi-conservative mutations with (.).

To interrogate binding interactions near this docking site, a single methionine (M141) residue

near the site was changed to an isoleucine, as shown in Figure 20:



Figure 20. Docking site 6. View of poliovirus 3Dpol from the top. Docking site 6 is marked in green, with the methionine 141 residue marked in magenta.

An NMR HSQC spectrum was then collected on this variant (Figure 21):



Figure 21. Comparison of WT and M141I variant 3D. The results of the two 500 MHz HSQC experiments are aligned. Superimposed 500 MHz HSQC spectra of poliovirus 3Dpol (200 uM) unbound, in black, and poliovirus 3Dpol in the presence of 2.7 mM PI4P, in magenta. These spectra were collected in 20 mM HEPES, pH 8.0, 150 mM NaCl, 1 mM DTT, and 0.02% sodium azide in D₂O.

The peaks at 2.25 ppm and 1.5 ppm in Figure 21 are PI4P. A lack of chemical shift changes in the relevant residues did not confirm the validity of docking solution 6. Although slight chemical shift changes were seen in some signals, these were not near the relevant M141. Many of these peaks are at different intensities, which makes it difficult to resolve all of the signals on the same spectrum. In particular, solvent exposed residues tend to show much higher signal intensities. As such, different parts of the spectra are shown separately in Figure 21.

Though none of the signals showing chemical shift or intensity changes were near docking site 6, a few of these signals were near docking sites 1 and 5. Specifically, M6 is found near docking site 1 and M392 and M394 are found near docking site 5 (Figure 22).



Figure 22. Top views of Poliovirus 3D. Docking site 1 is highlighted in magenta, docking site 5 is highlighted in blue, and docking site 6 is highlighted in red. Methionine residues nearby the docking sites are highlighted in orange.

At this point, the NMR titration experiments were consistent with docking sites 1 and 5 and did not show any evidence in favor of docking site 6. From these analyses, it appears that there may be an entire 3D surface across which a PI4P membrane could bind, perhaps in a multivalent fashion. Again, it should be noted that the 3D construct being used actually contained two amino acid substitutions (L446D, R455D) near docking site 5. Since these amino acid changes may influence the PI4P binding, a different anti-dimerization variant (D339A, S341A, D349A), developed by the Cameron group, was used for the following studies. Based on some positive results from binding assays performed by Dr. Craig Cameron's group, docking site 5 was chosen next for study (Figure 23).



Figure 23. Supported lipid bilayer binding assays for 3D variants. Data provided by Djoshkun Shengjuler. PI4P is shown to preferentially bind the true WT and D339A, S341A, D349A variant. In the L446D, R455D variant, the same trend is not seen.

Figure 23 was generated by using a supported lipid bilayer (SLB) assay. SLB allows for the modeling of membrane systems, which is useful in interrogating protein-PI4P interactions in a membrane-like environment. In particular, this assay was performed using planar SLBs in a microfluidic device to wash the protein over the ligand-containing ligand. Binding of the protein to the membrane was indicated by fluorescence due to the accumulation of hydroxide ions at the site of binding. This ion accumulation was due to the an increase in interfacial potential as the positively charged protein binds the membrane. Given the promise of the new anti-dimerization variant, an NMR HMQC titration experiment was performed (Figure 24):



Figure 24. Isoleucine-labeled spectrum of Poliovirus 3Dpol (D339A, S341A, D349A) with PI4P. Superimposed 600 MHz HMQC spectra of poliovirus 3D (200 uM) unbound, in black, and poliovirus 3D in the presence of 2.7 mM PI4P, in red. The buffer conditions were 25 mM potassium phosphate, pH 8.0, 150 mM NaCl, 1 mM DTT, and 0.02% sodium azide in D_2O .

While perhaps some very slight chemical shift changes were seen in some of the resonances, none

of them were at residues near any of the expected docking sites. While this result is interesting, it

necessitates further inquiry. A chemical shift change is seen in the I316 resonance. Though I316 is not

near any of the docking sites, it is near the three amino acids that showed chemical shift changes in Figure

17.

For comparison, the L446D, R455D variant apoprotein spectrum was overlaid with the D339,

S341, D349A apoprotein spectrum (Figure 25).



Figure 25. Overlay of **3D** variants. The 600 MHz HMQC of the two different **3D** variants tested are overlaid. The first variant is shown in black and the anti-dimerization variant is shown in blue. The spectra were aligned based on the most intense peak, at -0.15 in the proton dimension and 10.5 in the carbon dimension.

The two spectra were overlaid and the anti-dimerization variant was initially shifted upfield in the proton dimension and downfield in the carbon dimension. The spectra were then aligned based on the highest intensity peak. Interestingly, the anti-dimerization variant is still shifted downfield in the carbon dimension. This difference is likely due to a temperature difference, magnetic field change, and/or major conformation differences between the two variants. It is worthwhile to note that the 600 MHz NMR instrument was having problems holding temperatures around the time the anti-dimerization variant spectrum was collected.

The notable chemical shift changes in Figure 17 and Figure 24 were also plotted on the protein structure, to show relative positions of the residues (Figure 26).



Figure 26. Key protein residues in poliovirus 3D. Shown is a front view of the poliovirus 3D protein. Docking solution 1 is shown in magenta, docking solution 5 is shown in blue, and docking solution 6 is shown in red. Isoleucines that showed chemical shift changes in Figures 17 and 24 are highlighted in orange. Methionines that showed chemical shift changes in Figure 21 are highlighted in yellow.

Interestingly, it appears that most of the residues and the docking sites highlighted in Figure 26 are on the same face of the protein. Perhaps the 3D-PI4P interaction is actually a multivalent interaction that involves multiple docking sites on the same face of the protein. This sort of interaction at membranes is well documented throughout the literature, especially in relation to receptor-ligand binding and cell signaling.^{48,49} "Lipid rafts" in cellular membranes have also been implicated in multivalent protein

binding, especially when membranes are moving around the cytoplasm during signal transduction.⁵⁰ Similar interactions are also seen with peripheral membrane proteins used for a variety of cellular processes.⁵¹ As such, it seems reasonable that poliovirus 3D may be behaving in a similar manner, employing multivalent ligand binding to increase affinity or specificity.

Chapter 4:

Discussion

Docking Solutions 1 and 5

The goal of this study was to determine the validity of the docking solutions for PI4P to Poliovirus 3D binding using solution-state NMR. Based on the initial results (Figure 17), docking site 5 emerged as the most promising of the six docking solutions. In particular, the chemical shift changes seen in the I397 and I436 peaks indicated docking site 5 plays some role in PI4P binding. However, The chemical shift change associated with I58 is also important to note, as this residue is implicated in docking site 1. Interestingly, docking sites 1 and 5 do not fall within close proximity of each other, though they are on the same face of the protein.

That being said, docking site 5 showed further promise based on the work of Djoshkun Shengjuler in Dr. Craig Cameron's lab. Djoshkun performed lipid bilayer binding assays that showed an improved binding affinity for PI4P upon changing from the NMR "wild-type" (L446D, R455D) to a different anti-dimerization 3D variant (D339A, S341A, D349A).³² Because the NMR wild-type amino acid substitutions are near docking site 5 and the restoration of these residues to the natural wild-type restored PI4P binding, docking site 5 showed promise.

The effect of the dimerization amino acid changes also remained unclear. With the original L446D, R455D variant, binding of PI4P was in the millimolar range, which seems reasonable in comparison to the 3C protein, which showed binding constants in the millimolar range, as well.²¹ The uncertain nature of these results was furthered by the NMR results for the D339A, S341A, D349A variant, which showed virtually no signal changes upon addition of PI4P (Figure 24). All told, the NMR data was most consistent with docking sites 1 and 5.

Creation of a New Docking Site

Perhaps in an effort to create anti-dimerization variants useful for NMR study, additional docking site(s) were created. The exceptional initial promise of docking site 5 was rather curtailed with the D339A, S341A, D349A variant (Figure 24). While this spectrum does not preclude docking site 5 from truly being a site of PI4P interaction, it certainly does suggest there may be other factors at play.

It seems plausible that by way of creating these anti-polymerization mutants, additional docking sites are being created. However, the evidence for this idea is still slim, though it may warrant additional attention.

Future Directions

Effective Modeling of Membrane Interactions

As discussed earlier, the short-chain PI4P provided a qualitative analysis of the binding interactions. However, these studies modeled the interactions in a very artificial manner, as the free PI4P may adopt orientations in solution that it may not have liberty to adopt in a biological membrane. While the nanodiscs inch closer to a true membrane model, they also have shortcomings in issues of concentration, orientation, and curvature.

In recent years, technology has arisen that allows the study of membrane interactions in more native scenarios. In particular, studies have shown that solid-state NMR can be used to study supported lipid bilayer systems.⁵² In addition to nanodiscs, some groups have touted the use of lipid vesicles or bicelles of other morphologies to serve as membrane mimetics for NMR study.^{53,54,55}

Alternative Approaches

Perhaps the utility of HSQC/HMQC NMR in studying this interaction has gone as far as it can. The advantage of NMR is that it can provide site-specific information. As such, other types of NMR experiments may be useful in further elucidating 3D-PI4P interactions.

For example, spin relaxation experiments may be done, where a spin label is attached to the PI4P molecules. A loss of signal from other residues in such a spin relaxation experiment would indicate that those residues are interacting with the tagged PI4P. The binding of this altered PI4P could be confirmed using fluorescence polarization.

Additionally, should the binding site(s) be conclusively determined, other types of NMR experiments, like ZZ exchange, may be useful in studying the dynamics.

In the Context of 3CD

Perhaps the inconsistent nature of the results is indicative that the model used does not accurately reflect the *in vivo* system for the interaction of 3D and PI4P. Previous work showing that 3C binds PI4P exists.²¹ As such, it may be that the proposed PI4P localization mechanism acts on the 3CD molecule, before the individual 3D molecule has been cleaved away.

To further interrogate this notion, 3CD mutational analysis may be valuable. In the same way that 3D binding interactions were studied, so too may 3CD interactions be studied.

PI4P may also interact with the 3CD protein, considering that it is a crucial precursor protein of the viral protease (3C) and RNA-dependent RNA polymerase (3D)⁵⁶ that must be cleaved during post-translation polyprotein processing to activate 3D for genome replication. This 3CD polyprotein structurally resembles two balls on a string (Figure 27A), with a linker region connecting the 3D (Figure 27B) and 3C proteins (Figure 27C). Contained in this linker region is a recognition site that may be cleaved by 3C or 3CD molecules.^{57,58,59} Cleavage of the 3CD polyprotein into its component 3C and 3D

parts is imperative, as the 3D protein only displays polymerase activity once the cleavage has occurred.⁵⁹ 3CD may interact with PI4P through its 3C and/or 3D domains.



Figure 27. Crystal structure of 3CD and component 3C and 3D. A. The crystal structure of 3CD (PDB 2IJD) is shown with active site residues of 3C protease (H40, E71, C147) and 3D RdRp (D416, D511, D512) highlighted as red spheres. B. 3D subunit (PDB 1RA6) shown individually. C. 3C subunit (PDB 1L1N) shown individually.

In analyzing the 3D docking study data in conjunction with Dr. Chan's 3C data²¹ the

possibility of a 3CD multivalent docking surface emerged (Figure 28).



Figure 28. 3CD docking sites. Residues showing chemical shift changes in 3C and 3D are highlighted in the context of 3CD.

The relative locations of the residues showing chemical shift changes perhaps indicate

that the 3CD molecule may be localizing to intracellular host membranes by docking to PI4P

through a multivalent interaction with sites on both the 3C and 3D domains.

BIBLIOGRAPHY

1. Breitbart, M and Rohwer, F. (2005) Here a virus, there a virus, everywhere the same virus? *Curr*. *Trends Microbiol.*, 13, 278-284.

2. Edwards, R. and Rohwer F. (2005). Viral metagenomics, Nature Rev Microbiol., 3, 801-805.

3. Ahlquist, P. (2006) Parallels among positive-strand RNA viruses , reverse-transcribing viruses and double-stranded RNA viruses. *Nature Rev. Microbiol.*, 4, 371–382.

4. Carter, J and Saunders, V. (2013) Virology: Principles and applications. Wiley.

5. Ferrer-orta, C., Arias, A., Escarmi, C. & Verdaguer, N. (2006) A comparison of viral RNA-dependent RNA polymerases. *Curr. Opin. Struct. Biol.*, 16, 27–34.

 Koonin, EV. (1991) Genome Replication/Expression Strategies of Positive-Strand RNA Viruses : A Simple Version of a Combinatorial Classification and Prediction of New Strategies. *Virus Genes*, 70, 273–281.

Koonin, EV, Senkevich, TG, Dolja, VV. (2006) The ancient Virus World and evolution of cells.
 Biology Direct, 29, 1–27.

8. Hayes, RJ and Buck, KW. (1990) Complete Replication of a Eukaryotic In Vitro by a Purified RNA-Dependent RNA Polymerase Virus RNA. *Cell*, 63, 363–368.

9. Ryu, WS. (2016) Molecular Virology of Human Pathogenic Viruses. Academic Press, Korea, 153-164.

10. Behrens, S, Tomei, L, De Francesco, R. (1996) Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.*, 15, 12–22.

11.Ferrer-orta, C, Ferrero, D, Verdaguer, N. (2015) RNA-Dependent RNA Polymerases of

Picornaviruses: From the Structure to Regulatory Mechanisms. Viruses, 7, 4438-4460.

12. Kopek, BG, Perkins, G, Miller, DJ, Ellisman, MH, Ahlquist, P. (2007) Three-Dimensional Analysis of a Viral RNA Replication Complex Reveals a virus-induced mini-organelle. *PLoS Biol.*, 5, 2022-2034.

13. van der Hoeven, B, et al. (2016) Biogenesis and architecture of arterivirus replication organelles. *Virus Research*, 220, 70-90.

14. Belov, GA and van Kuppeveld, FJM. (2012) (+) RNA viruses rewire cellular pathways to build replication organelles. *Curr. Opin. Virol*, 2, 740–747.

15. Hsu, N. et al. (2010) Viral Reorganization of the Secretory Pathway Generates Distinct Organelles for RNA Replication. *Cell* 141, 799–811.

16. Den Boon, J, Ahlquist, P. (2010) Organelle-Like Membrane Compartmentalization of Positive-Strand RNA Virus Replication Factories. *Annu. Rev. Microbiol.*, 64, 241-256.

17. Belov, GA, et al. (2011) Complex dynamic development of poliovirus membranous replication complexes. *J. Virol.*, 86, 302-312.

18. Blanchard, E and Roinegeard, P. (2015) Virus-induced double-membrane vesicles. *Cell. Microbiol.*, 17, 45-50.

19. Limpens, RWAL, et al. (2011) The transformation of enterovirus replication structures: A threedimensional study of single- and double-membrane compartments. *mBio*, 2, 1-10.

20. Belov, GA. (2014) Modulation of lipid synthesis and trafficking pathways by picornaviruses. *Curr. Opin. Virol.*, 9, 19–23.

21. Chan, YM. (2015) Functional regulation and protein structural dynamics of poliovirus 3C protease (Doctoral dissertation).

22. Di Paolo, G, De Camilli, P. (2006) Phosphoinositides in cell regulation and membrane dynamics. *Nature.*, 443, 651-657.

23. Bishé, B, Syed, GH, Field, SJ, Siddiqui, A. (2012) Role of Phosphatidylinositol 4-Phosphate (PI4P) and its Binding Protein GOLPH3 in Hepatitis C Virus Secretion. *J. Biol. Chem.*, 287, 27637–27647.
24. Arita, M. (2016) Mechanism of Poliovirus Resistance to Host Phosphatidylinositol-4 Kinase III β
Inhibitor. *ACS Infect. Dis.*, 2, 140-148.

25. Hansen, JL, Long, AM, Schultz, SC. (1997) Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure*, 5, 1109–1122.

26. Pfeiffer, JK and Kirkegaard, KA. (2003) A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc. Natl. Acad. Sci. U.S.A.*, 100, 7289–7294.

27. Thompson, AA and Peersen, OB. (20014) Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent. *EMBO J.*, 23, 3462–3471.

28. Gohara, D. W. et al. (2000) Poliovirus RNA-dependent RNA Polymerase (3D pol). *J. Biol. Chem.*, 275, 25523–25532.

29. Ng, K, Arnold, JJ, Cameron, CE. (2008) Structure-function relationships among RNA-dependent RNA polymerases. *Curr. Top Microbiol. Immunol.*, 320, 137–156.

30. Pata, JD, Schultz, SC, Kirkegaard, K. (1995) Functional oligomerization of poliovirus RNAdependent RNA polymerase. *RNA*, 1, 466–477.

31. Yang, X. et al. (2012) Article Motif D of Viral RNA-Dependent RNA Polymerases Determines Efficiency and Fidelity of Nucleotide Addition. *Struct. Des.*, 20, 1519–1527.

32. Shengjuler, D, et al. (2017) Poliovirus 3C protein is a phosphoinositide-binding protein (Unpublished).

33. Morris, GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, and Olson AJ. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. *J. Comput. Chem.*, 16, 2785-2791.

34. Dolinsky, TJ, Czodrowski, P, Li, H, Nielsen, JE, Jensen, JH, Klebe, G, Baker, NA. (2007) PDB2PQR:
Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucl. Acids Res.*, 35, 522-525.

35. Dolinsky, TJ, Nielsen, JE, McCammon, JA, Baker, NA. (2004) PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucl. Acids Res.*, 32, 665-667.

36. Denk, M. (2005) CHEM 2070 structure and spectroscopy. Online.

37. Wiesner, S and Sprangers, R. (2015). Methyl groups as NMR probes for biomolecular interactions. *Curr. Opin. Struct. Biol.*, 35, 60–67.

38. Tugarinov, V, Kanelis, V, Kay, LE. (2006) Isotope labeling strategies for the study of highmolecular-weight proteins by solution NMR spectroscopy, *Nature Protocols*, 1, 749–754.

39. Ayala, I, Hamelin, O, Amero, C, Pessey, O, Plevin, MJ. (2012) An optimized isotopic labeling strategy of isoleucine- γ_2 methyl groups for solution NMR studies of high molecular weight proteins. *Chem. Commun.*, 48, 1434–1436.

40. Fasshuber, KH, Demers, J, Chevelkov, V, Giller, K, Becker, S, Lange, A. (2015) Specific ¹³C labeling of leucine , valine and isoleucine methyl groups for unambiguous detection of long-range restraints in protein solid-state NMR studies. *Journal of Magnetic Resonance*, 252, 10–19.

41. Tugarinow, V., Kay L.E. (2008). Sigma Aldrich. ISOTEC stable isotopes. Online.

42. Mitra, N. (2013) Nanodiscs: Membrane protein research in near-native conditions. *Mater. Methods*, 3, 177.

43. Liu Y, Kahn RA, Prestegard JH. (2014). Interaction of FAPP1 with ARF1 and PI4P at a membrane surface: an example of coincidence detection. *Structure*, 22, 421-430.

44. Ritchie, TK, et al. (2014). Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol*, 464, 211-31.

45. He, J. et al. (2011) Molecular Basis of Phosphatidylinositol 4-Phosphate and ARF1 GTPase Recognition by the FAPP1 Pleckstrin Homology. *J. Biol. Chem.*, 286, 18650–18657. 46. Kobashigawa Y, Harada K, Yoshida N, Ogura K, Inagaki F. (2011) Phosphoinositide-incorporated lipid-protein nanodiscs: A tool for studying protein-lipid interactions. *Analytical Biochemistry*, 410, 77-83.

47. Cavanagh, J, Fairbrother, W, Palmer III, AG, Skleton, NJ. Protein NMR spectroscopy – principles and practice. *Academic Press*.

48. Zheng, H and Du, X. (2011) Multivalent protein binding in carbohydrate-functionalized monolayers through protein-directed rearrangement and reorientation of glycolipids at the air-water interface. *Biochim. Biophys. Acta.*, 1808, 2128-2135.

49. Jung, H, et al. (2010) Multivalent ligand-receptor binding on supported lipid bilayers. *J. Struct. Biol.*, 168, 90-94.

50. Simons, K and Ikonen, E. (1997) Functional rafts in cell membranes. Nature, 387, 569-572.

51. Cho, W and Stahelin RV. (2005) Membrane-protein interactions in cell signaling and membrane trafficking. *Annu. Rev. Biophys.*, 34, 119-151.

52. Salnikov, ES, et al. (2015) Solid-state NMR/dynamic nuclear polarization of polypeptides in planar supported lipid bilayers. *J. Phys. Chem. B.*, 119, 14574-14583.

53. Sanders, CR., Oxenoid, K. (2000) Customizing model membranes and samples for NMR spectroscopic studies of complex membrane proteins. *Biochim. Biophys. Acta*, 1508, 129-145.

54. Warschawski, DE, et al. (2011) Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochim. Biophys. Acta.*, 1808, 1957-74.

55. Mäler, L. (2011) Solution NMR studies of peptide-lipid interactions in model membranes. *Molecular Membrane Biology*, 29, 155-176.

56. Marcotte, LL. et al. (2007) Crystal Structure of Poliovirus 3CD Protein: Virally Encoded Protease and Precursor to the RNA-Dependent RNA Polymerase. *Journal of Virology*, 81, 3583–3596.

57. Blair, WS, Nguyen, JHC, Parsley, TB, Semler, BL. (1996) Mutations in the Poliovirus 3CD

Proteinase S1-Specificity Pocket Affect Substrate Recognition and RNA Binding. Virology, 218, 1–13.

58. Cornell, CT and Semler, BL. (2002) Subdomain Specific Functions of the RNA Polymerase Region of Poliovirus 3CD Polypeptide. *Virology*, 298, 200–213.

59. Harris, KS, Reddigari, SR, Nicklin, MJH, Hammerle, T, Wimmer, E. (1992) Purification and

Characterization of Poliovirus Polypeptide 3CD, a Proteinase and a Precursor for RNA Polymerase.

Journal of Virology, 66, 7481–7489.

ACADEMIC VITA

Rohan Basu rohan.basu1010@gmail.com

Education • BS in Biochemistry Minor: Chemistry **Research Experience and Awards** Researcher at The Pennsylvania State University, David Boehr Lab 2015-Present Wrote honors thesis on research **Biological Sciences Trainee, Centers for Disease Control and Prevention** 2016 Co-authored conference presentation Designed and setup up experiments to test effectiveness of respiratory personal protective equipment • Began drafting and editing manuscript for publication • Created and delivered presentation to present to entire research branch • Summer Researcher at University of Pittsburgh, Angela Gronenborn Lab 2014-2015 Published research in Journal of Biological Chemistry: DOI:10.1074/jbc.M116.740415 Wrote and delivered a public talk on research • Researcher at The Pennsylvania State University, Song Tan Lab 2014-2015 Wrote successful undergraduate grant application to university • Created and presented multiple posters detailing performed research • Henrietta M. Fisher Memorial Award 2016-2017 For high-achieving students from Dauphin County • **Teaching and Work Experience** Private Instructor, PSU Knowhow/Gradepeak 2015-Present Worked one-on-one to improve students' understanding of science subjects 2016-Present Short-Term Crisis Counselor, Community Help Centre • Counsel and assist clients with basic needs, mental health, and day-to-day problems Underwent 180 hours of intensive training to assist State College community as a short-term crisis counselor Participate in staff-run agency decisions • **DUI Instructor, Centre County Correctional Facility** 2017-Present Teach drug and alcohol education class to DUI offenders at local correctional facility **Organizer, Soles4Soles** 2014 Created, organized, promoted, and hosted school-wide shoe drive for underprivileged people THON, Lambda Chi Alpha Fraternity 2015 Developed and maintained close relationships with organization sponsored families

 Advertised and solicited donations for Penn State Dance Marathon (THON) to benefit families affected by pediatric cancer • Organized two new fundraisers to support families affected by pediatric cancer and fund research

Other Volunteering Activities

• Member of communications committee for Penn State IFC/Panhellenic Dance Marathon (THON) 2013-14

2016-Present

• Mt. Nittany Medical Center, Volunteer